The saguenay population register and the processing of occupational data: an overview of the methodology

Bouchard, Gerard

THE SAGUENAY POPULATION REGISTER AND THE PROCESSING OF OCCUPATIONAL DATA: AN OVERVIEW OF THE METHODOLOGY(*)

Gérard Bouchard (+)

Abstract: This paper outlines the research that has been carried out in the last few years within a Canadian research group (SOREP) to device a whole new methodology dealing with occupational data. These data come from a regional, computerized population register and they have been submitted to several processings in order to bring out their significance and possible uses. The study reveals that the social inquiries - namely into historical mobility - can be dramatically biased by a lack of preliminary, critical work. It also provides a new occupational classification grid which, hopefully, will avoid some of the usual, major known pitfalls. Finally, the paper adresses the problem of assessing the content of occupational titles, which is a necessary step allowing a proper classification into categories. Here, a basic methodological problem is pointed out: the criteria and guidelines that should be retained in order to achieve a consistent and systematic classifying process. It is shown that the use of contemporaneous technical scales, even to classify past occupational data, doesn't necessarily entails anachronism.

Our population register now covers the entire Saguenay region for the period 1842 - 1971. Begun in 1972, it consists exclusively of name-data, drawn premarily from parish registers, from which 650 000 certificates have been transcribed and computerized. These data, processed through linkage programmes have provided enormous quantities of occupational information. Automatic family reconstitution allows this information to be organized into individual and family biographies, which in turn facilitate the construction of long-term occupational itineraries. Used in this way, occupational data can support enquiries into social mobility, but their usefulness goes further, because they also feed all those studies seeking to identify the distribution of a specific trait across a population, to characterize one group in relation to another, to provide samples, to identify the individuals in a community, etc. For the general purposes of the use of the register, in very different fields of knowledge and perspectives, it seemed worthwhile to develop procedures and instruments to insure appopriate use of the occupational data, meeting the three following conditions: a) to safeguard the information as much as possible, b) to avoid changing it in any way, particularly in the process of classifying and grouping, and c) to reduce, as little as possible, the scope of the uses to which it can be put.

We present those procedures and instruments here, no doubt too briefly, aware of how precarious are some of the choices and how many problems remain to be solved - the difficulties which typify this field of research being, as we all know, a school for modesty. Our method consists of a six-point approach:

(+) Address all communications to: Gérard Bouchard, Université du Québec à Chicoutime, Societé de Recherche Sur Les Populations, Chicoutimi, Canada.

1. A critical examination of the nature of the occupations data and their significance.
2. The construction of categories, in the form of an occupational classification grid which organizes the occupations into only descriptive groups, which are not arranged in terms of social hierarchy.
3. The problem of classifying the occupations, insofar as this involves a knowledge and a rigorous evaluation of their content.
4. The difficulties raised by this evaluation in longitudinal studies, in that technological changes continually threaten to change the content or parameters of an occupation, and so invalidate the evaluation made of it.
5. The use of occupational categories in the analyses of social hierarchies.
6. The classification of occupations according to sectors of economic activity.

It must be remembered that the classifations discussed are all reversible. Also, the survey and data-collecting were executed so as to preserve access to the original data, which are thus available for other kinds of classification. More generally, it is also very important to state that this overall methodology has been defined with reference to a very specific purpose, that is, to the exploitation of a large data base which is a public and a pluridisciplinary resource. Therefore it has to be available to a very large array of users and scientific goals. We believed that this consideration ruled out the possibility of grounding our methodology on a very particualr and limited theorical orientation hence the empirical style of this approach. However, for those users who would like to rely on different methods, it is always possible to access the raw data and process them in their own ways.

1. THE NATURE OF OCCUPATIONAL DATA

Before making any use of information about occupations, we must understand their content. In the context of studies of social mobility, for example, the problem might be illustrated as follows: if we know that Georges Tremblay, who lived at Jonquière from 1860 to 1929, exercised occupation "X" in 1895, one might infer the nature of his work, the kind of life he lived and, perhaps, his social position. And if, in 1920, the same Georges Tremblay declared himself to be of the occupation "Y", we would then deduce what he had become and would perhaps run the risk of making an evaluation, in terms of mobility, from "X" to "Y". But should an unexpected source suddenly reveal 15 or 20 other occupations declared by this individual between 1881 and 1927, far from clarifying the situation, everything could get confused: "X" and "Y" could now seem to be synonymous or equivalent titles; they could be spliced into a sequence of four or five different titles making up a new itinerary; this sequence might give place not only to changes but also to the parallel exercise of different occupations or double occupations. In short, the wealth of the data reveals the traps and idiosyncrasies of occupational language, and a knowledge of this grammar is, of course, a precondition of all analysis.

In the population register, we constructed a sample of 88,000 certificates of birth, marriage and death from six urban and rural Saguenay parishes. Using family reconstitions programmes, we obtained a file of 27,901 biographies in the form of occupational itineraries. Broken down into their constituent sequences, these itineraries were subjected to close analyses, which allowed us first to recognize the content of the occupation titles and then to construct a typology for occupational language. The nine following types of mentions of titles could then be identified:

1. Fields or sectors of activity
   Examples: Public employee, civil servant
   Forestry worker
   Construction worker
   Navigator
   Business man

2. Types or fields of training
   Examples: Musicians
   Geographer
   Economist
   Theologian
   Technician
   Engineer

3. Categories or groupings of trades or occupations
   Examples: Craftsman
   Artist
   Journeyman
   Labourer
   Agent

4. Civic or public duties
   Examples: Organist
   Town clerk in a small municipality
   Mayor
   Alderman
   County chief
   Bandmaster

5. Honorary ranks and titles
   Examples: Canon
   Director
   Squire
   Count

6. Administrative and institutional conditions or statuses
   Examples: Religions, priest
   Pensioner, fund-holder, retired
   Student
   Soldier
   Unemployed

7. Technical roles in industrial trades, related to the division of labour
   Examples: Baker
   Tester
   Stationman
   Sectionman
   Riveter
   Molder

8. Trades, occupations as such
   Examples: Shoemaker
   Blacksmith
   Book-keeper
   Judge
   Carpenter
   Farmer

Teacher
Carter

9. Double Occupations
Examples: Farmer/Seaman
Merchant/Shoemaker
Carpenter/Businessman
Bailiff/Forest ranger

This typology emphasizes the highly heterogeneous nature of occupational data, which is a serious obstacle to comparative studies and studies in social mobility. In this last case, for example, the occupational itineraries extracted from family histories reveal many pairs or sequences of occupations, where "A" - "B" (a succession of two different titles within one itinerary). Some of these sequences are real changes, others not. The task of deciding which is which is both the most difficult and the most essential, if we are to avoid a major overevaluation of mobility. To that end, we have produced a computer list of all the pairs with their frequencies, as well as a list of the cycles, these being defined as sequences of three occupations, of which the first and third are identical ("A" - "B" - "A"). The analysis of these data allowed us to separate with certainly the false changes, the near-changes and the real changes.

A) The false changes.

These take three forms:

1. The errors
These occur when the original information is registered by the priest, or when the data is collected or at the time of input at the terminal. Comparison of the codes allows these to be identified easily, especially in the case of the errors of collection and data entry, the latter being most of the time due to the confusion of two consecutive or similar numbers (3, 5, 8 ...). For example:

132 - Farmer / 032 Surveyor
132 - Farmer / 232 Real Estate Agent
042 - Lawyer / 242 Computer Scientist
256 - Journeyman / 286 Housewife

Of 1,074 different pairs examined, the file contained 14 errors of this type, or 1.3% of the total.

2. The non-occupations
Institutional and administrative status, public duties and responsibilities and official titles in hierarchies cannot be considered to be occupations. They appear in 18 sequences (1.7%).

3. Synonyms
First, there are synonyms in the strict sense, like police officer/constable, doctor/physician, etc. Their number is insignificant. Most frequent are titles which became equivalent in a particular historical context. Examples are:

Goldsmith / Jeweller
Hunter / Trapper
Carpenter / Cabinet-maker
Insurance agent/ broker

Office clerk / Book-keeper
Mason / Stonecutter
Plumber / Tinsmith
Journeyman / Labourer
Etc.

126 cases of synonymy were collected, involving 11.7% of the sequences, which gives a total of 158 false changes or 14.7% of the whole.

B) Near-changes

Here also, there are three kinds:

1. Pairs linking an occupation with a category
In all cases where the occupation involved is very obviously contained in the category, we cannot really speak of changes. Some examples are:

Journeyman / Baker
" / Boatman
" / Verger
" / Caulker
" / Roadmender
" / Carder
" / Carter
Artisan / Bricklayer
/ Cheesemaker
" / Carpenter
Labourer / Carter
" / Wheelwright
" / Clerk
" / Joiner

The case of the journeyman is, of course, the most interesting one. It is very obvious that this title indicates, not a craft, but the absence of craft. The journeyman is unstable by definition. He is the one who, lacking a particular trade, must move from one to another, usually to jobs which require little training.

The fields and sectors of activity, and the types of training and competence have been treated like the categories. Overall, these titles occur in 134 pairs (12.5%).

2. Pairs linking derived or parallel occupations
Some pairs link similar or close titles, relating two occupations which can be held simultaneously, or two jobs which were poorly differentiated in daily practice. They are usually revealed by studies of cycles and of occupations registered at short intervals (e.g. 1- 30 days). They cannot be called changes in the strict sense:

Farmer / Cheesemaker
Saddler / Shoemaker
Carpenter / Wheelwright
Tradesman / Bottler
Metal-caster / Blacksmith
Nurse / Hygienist
Inspector / Tax-collector
Bookseller / Bookbinder
Mechanic / Welder
Merchant / Wholesaler

We counted 86 pairs of this type, representing 8.0% of the whole.

3. Pairs linking an occupation and a double occupation
This case closely resembles the preceding one and can doubtless be attributed to the same causes:

Baker / Baker-tradesman
Farm lobourer / Farm labourer-seaman
Notary / Court Clerk-Notary
etc.

There were 20 pairs of this type (1.9%). The total number of near-changes is thus 240, or 22.3% of the pairs.

To sum up, of 1,074 pairs or presumed changes, only 676, or 63.0% can be considered genuine (see Table 1). On the other hand, of the 231 occupational titles appearing on the sample file between 1842 and 1941, 78 should be eradicated from the list (synonyms, non-occupations) and 16 should be regarded as suspect (categories and other assimilated titles). This leaves 137 occupations titles, or 59.3% of the total, which can be entered without fear of causing error in the analysis of the changes.(1)

In themselves, these last two figures show quite clearly how very fluid occupational data are, and how necessary it is that any use of them be preceded, as we said, by careful examination.(2) Of course, we don't believe that the equivalencies that we have established in the Saguenay case should be the same elsewhere. It must be clear that this work has to be made within every local historical context. Also, it is worth mentioning that the data considered as synonyms or false and near-changes are not discarded; but their use is subject to appropriate treatment and controls.

2. THE CONSTRUCTION OF OCCUPATIONAL CATEGORIES

Once verified and cleaned, the occupational data can be classified in three ways, which must be carefully defined at the outset. The first type of classification sorts the individuals and occupations according to sectors of economic activity. Avoiding all forms of ranking, it allows to know the structure of the manpower at all times, and to trace its evolution. The division used thus refer to the fields of production (retail trade, construction, etc.). A second type of classification is based on occupational categories. Here, the occupations are grouped on the basis of similarities inherent to the content of the tasks themselves. The internal, technical dimension of the work performed is therefore the dominant element in the definition of the categories, which are strictly descriptive groups, not ranked by the usual social criteria (e.g. administrators, professionals, craftsmen, etc.). Finally, a third type is designed to establish social rankings by creating an order, either among the categories, or among the occupations considered on an individual basis. The scales used are strata or classes, or all forms of social position; they are principally defined by social criteria like income or wealth, prestige, lifestyle, power, etc.

In this section, we present a new grid of occupational classification, used to deal with Saguenay data. We do not think it possible to avoid this recourse to categories, as the approximately 1,200 occupational titles in

TABLE 1

Statistics of false, true and near-changes in terms of the number of pairs of occupation titles. Saguenay, 1842-1941.
(Source of data: parish records)

TYPES OF PAIRS	NUMBERS	%
1- False changes	158	14.7%
- Errors	14	1.3
- Non-occupations	18	1.7
- Synonyms	126	11.7
2- Near-changes	240	22.4
- Categories	134	12.5
- Derived occupations	86	8.0
- Double occupations	20	1.9
3- True changes	676	63.0
TOTAL	1074	100.0

(SOREP) Saguenay Social History Research Project.

the file can hardly be treated individually. Furthermore, this move to categorization makes it easier to overcome difficulties resulting from the redundant and very fragile nature of occupational titles, as we saw, since synonyms and near-synonyms often fall into the same category. Thirdly, the study of the vertical mobility within the occupations themselves raises many as yet unresolved methodological problems.(3) Finally, the study of social hierarchies is intended to represent a number of strata, classes or rankds corresponding to levels of power, wealth, prestige, etc. But these remain abstract entities as long as their content has not been identified: who are the individuals in such positions or in such strata? Which elements make up the upper and middle classes? Or the working class? Inevitably, these attempts at identification rely on the occupations or, more likely, on the categories.

A) Notes on the methodology of occupational classification

Our research group tried out various classification grids before arriving at the idea of a new instrument. This preliminary step allowed us to identify what seems to be three forms of methodological deficiency.(4)

1. Confusion of categories with hierarchies
   Many instruments consider as occupational categories groupings which are actually strata, classes or hierarchic positions. This is the case, for example, with all scales using divisions such as Gentlemen or Proprietor, which gather individuals who could belong to very different occupations and categories. This raises the problem of the choice of the criteria used to define the categories. It is obvious that ownership amd nobility are criteria relating primarily to social hierarchy in the strict sense, and are inappropriate for the classification of occupations.

2. Redundant categories
   This second, very frequent problem, follows from the first: the categories are often so constructed that many occupations could receive more than one classification. This is the case when one category overlaps another, or even contains it entirely. For example:

   a - White-collar worker
   b - Manual Lobourer
   c - Service sector
   d - Primary sector

   The primary sector clearly contains both manual and nonmanual occupations. On this basis, how should we classify mining engineers, woodcutters or farm-labourers? The Service sector category raises similar problems; they affect a considerable number of classification grids. Here is another example of the same kind:

   a - Agriculture
   b - Industry
   c - Commerce
   d - Trade-related occupations
   e - Professionals
   f - Retired (fund-holders)
   g - Proprietors (ouners)
   h - Unemployed

Within this list of categories (which the author calls "socio-occupational groups"), only two refer to groups of related occupations (d and e). The others refer either to sectors of economic activity (a, b, and c), or to socio-economic status. Under these conditions, one wonders how the classification can be performed with proper respect for the uniformity which is so essential in this matter.

3. Imprecise criteria for definition and classification
All the above seems to be largely the result of lack of attention and thoroughness in the choice of criteria for defining categories and for classifying the occupations.
In the first place, this can lead to the use of social classes or single occupations, or both at once, as occupational categories. Second, and most frequent, is the error of inserting several unannounced criteria into the construction of the categories; or, having announced them clearly, of omitting to establish an order of precedence governing their application, thus giving rise to a further source of incoherence.

These reflections urge a sharper awareness in the making of occupational classifications.(5) We have restricted ourselves here to an examination of the logical structure of the instruments. One can imaging what a similar exercise might reveal when applied to the results of the analyses - but this would undoubtedly require laborious and, perhaps, impossible reconstructions. We were convinced, from this, of the value of attempting to construct a new grid for occupational categories, based on explicit criteria and rules, and aimed at overcoming to the greatest possible extent of difficulties just indicated.

B) A new classification grid

Again, it is important to draw a careful distinction between occupational categories and social hierarchies. The concept of an occupation refers, in the first place, to an internal and endogenous dimension, which is the work process in itself, or the whole of the actions, conditions and material or technical components inherent to the execution of the task. In the second place, it refers to an external, or exogenous dimension, which is the social insertion or extensions of the activity of the work. Parallel to a technical, internal order, the occupation is defined socially in terms of levels and ways of life, of solidarities and social relationships. In this perspective, the categories are descriptive entities, referring only to the technical organization and immediate content of the jobs. The purpose of the hierarchies is to establish an order among the categories, an order which reflects what is usually called the social division of labour; that is, the economic, cultural and social attributes associated with an occupation or a set of occupations. According to our procedure, the construction of the categories is a separate step preceding the construction of the hierarchies, and it should be drawn on criteria specific to it.(6)

1. Preliminary rules
The distinction just established suggests a first rule:

Rule No. 1: To refer, in the classification of an occupation, only to criteria relating to internal or technical dimensions of that occupation, and to regroup the titles on the basis of intrinsic similarities between the activities of the work.

Moreover, as far as the problems of redundancy are concerned, and in the case of overlaps between excessively broad categories, or those with very mixed content, it is generally obvious that the criteria for grouping are badly defined. Hence the second rule:

Rule No. 2: To adopt from the start a number of explicitly defined criteria, each of which refers to a specific aspect of the activity involved in the job.

These precautions do not ensure appropriate classification; it is still necessary that the criteria adopted be applied uniformly and always in the same order. Actually, it frequently occurs that two sets of criteria will not produce the same classification, depending on the order in which they are applied to an occupation. This leads to a third rule:

Rule No. 3: To establish an order of precedence or priority among the criteria, which will control their application in the classification process.

Finally, two further rules provide for a flexible and more effective use of the grid. One relates to number of categories:

Rule No. 4: To construct categories numerous enough a) to obtain very homogenous sub-sets, close to raw data, and b) to make it possible to aggregate categories, according to the various need of the users.

Given the imprecision of occupational language, we designed the final rule to minimize the amount of information lost or altered during classification:

Rule No. 5: To provide categories which can admit occupational titles containing some elements of imprecision, in order to avoid either excessive use of the category "unclassified" (loss of information) or erroneous classifications (alteration of the information).

It seems to us that the application of these five rules must ensure sufficient cohesiveness in the classification process. We can also expect that the resultant groupings will satisfy the need for thoroughness, since each occupation must be assigned a specific category. The success of the procedure then depends on the definition of appropriate criteria for classification.

C) The Classification criteria (7)

Two constraints governed the choice of the five criteria used; a) their aptitude for the creation of useful and familiar categories; b) their potential for easy use, taking into account what we know of the occupations to be classified - we must remember that parish register certificates only give the title of the occupation, and no more. This being said, we recognize, of course, that the direction we have taken is just one among many others, probably as relevant.

We have selected five traits or characteristics from the principal components of the workplace, which are used as criteria defining occupational categories. There are, first, two basic criteria, applicable to all

TABLE 2

GRID FOR CLASSIFYING OCCUPATIONS

BASIC CRITERIA			AUXILIARY CRITERIA			NOMENCLATURE OF THE CATEGORIES (in order of the matrix)
I- NATURE OF THE EFFORT	COMPLEXITY: II- Technical difficulty	COMPLEXITY: III- Responsibility	IV- Area of impact of the management unit.	V- Legal status of the management unit.	VI- Economic sector into which the task falls.	
A) Non manual	A) High level	A) High level	A) Supra-municipal, regional or extra-regional	A) Private		1. Directors of large businesses
				B) Public or para-public		2. Senior Officials
				C) Unknown		3. Semi-unknown (1-2)
			B) Municipal	D) Private		4. Small businessmen and industrials
				E) Public or para-public		5. Local Officials
				F) Unknown		6. Semi-unknown (4-5)
			C) Unknown	G) Private	A) Manufacturing and sales	7. Businessmen (1-4)
					B) Professional services	8. Professionals ("liberal prof.")
				H) Public or para-public		9. Public Administrators
				I) Unknown		10. Semi-unknown (7-8-9)
		B) Intermediate level				11. Middle managers
		C) Low level				12. Scientists, researchers, specialized white-collar workers
		D) Level unknown				13. Semi-unknown (1-12)
	B) Low level					14. Clerical workers, semi and non-skilled white collar
	C) Unknown					15. Semi-unknown (1-14)
B) Manual	D) High level	E) High level			C) Production of raw materials	16. Farmers, stock-breeders and assimilated trades
					D) Production of manufactured goods	17. Craftsmen (artisans)
					E) Unknown	18. Semi-unknown (16-17)
		F) Low level				19. Skilled workers
		G) Level unknown				20. Tradesmen (16)
	E) Low level					21. Semi and non-skilled workers
	F) Unknown					22. Manual workers (16-19)
C) Unknown						23. Unknown (1-22)
D) Unclassified						24. Others.

occupations, in the following order:

a) The nature of the effort required by the task or job.
This criterion consists in evaluating the intensity of physical effort and trying to recognize manual or non-manual dominance. It makes an initial separation of occupations, based on this elementary division.

b) The complexity of the task.
This dimension refers to the total weigh involved in executing the task. It is viewed in terms of two indicators, which are the level of responsibility or management (degree of control over one's own task and over those of others) and the technical difficulty associated with each occupation.
Three auxiliary criteria are then applied: the area of impact of the management unit (establishment-employer), its legal status (private, public, parapublic) and its field of activity.

In this way, we obtain a grid of twenty-four occupational categories, of which eight are allotted to titles including some sort of unknown (see Table 2). We cannot insist too much on the usefulness of this last provision. For example, in the case of a mason or a bricklayer, we do not know the level of independence but we do know that he is a manual labourer. Likewise, with an insurance salesman, we do not know the area of impact, but we do know the nature of the effort (non-manual). Those are the elements of information that we wished to preserve by creating the eight categories labelled semi-unknown.

3. EVALUATION OF THE CONTENT OF OCCUPATION

So, the grid provides a list of categories derived from a certain number of criteria. But the classification of occupations required an exact evaluation of their content, measured against these criteria. This means evaluating, for each occupational title, the nature of the effort, the level of difficulty, the degree of control, etc. The enormous difficulty of this task cannot be overemplasized (one only has to think of the number of titles involved), nor can we omit the thorough knowledge of the workplace needed, especially if we are resorting to quite refined categories.(8)

Obviously this knowledge is not available in social history, and cannot, perhaps, be acquired at all. What, then, are the methods of evaluation in current use? In fact, most authors neglect to inform their readers about this. All we can say is that they are often implicit or intuitive and vary considerably from one researcher to another, which hardly contributes to the clarification of the scientific debate.

Faced with this problem, our team first developed a tedious but apparently reliable process, based ultimately on the collective judgement of the researchers.(9) Nevertheless, we had to abandon it after a year, because we were unable to make decisions with uniformity and rigor. Here are some examples of situations where it seems the arbitrary does prevail over coherence, the decision being as likely to go in one direction as the other.

Classification by nature of the effort (Manual/non-manual)

Policeman	Airplane pilot
Watchman	Restaurant chief
Grocery-store manager	Tester
Computer operator	Maintenance supervisor

Classification by degree of difficulty (specialized/semi-/non specialized)

Truck driver	Butcher
Book-keeper	Seaman
Carter	Inspector
Hairdresser	Railway ticket agent
Bottler	Trapper
Barman	

The accumulation of these cases (and of incoherent decisions) induced us to change methods and seek elements of evaluation which would give the decision-making the necessary rigour and uniformity. From now on, when the title alone is not enough for the evaluation, we have recourse on some occasions to the semi-unknown categories. This is the case with the criterion of impact or the degree of responsibility.(10) For two other criteria, however, the nature of the effort and the level of difficulty, we believe we have found an adequate instrument in a descriptive list of occupations, produced by the Department of Manpower of the government of Canada (cf. the Canadian Directory of Classifications and Occupations: hereinafter CDCO). This directory, analogous to the American Dictionnary of Occupational Titles, necessitated the analysis of 20,000 occupational titles. For each of the 6 700 titles kept, it provides a very detailed description of the content of the task. This description takes note of eight parameters or dimensions (data, persons, things, specific occupational preparation, physical activities, etc.). Each of them is measured on a scale on which the occupation is situated by means of an index. By combining these dimensions and indices and matching them in relation to appropriate thresholds, we have developed a code which allows us to establish in a coherent and uniform way the level of difficulty appropriate to each occupation, and whether it is predominantly manual or non-manual.

This method has been described elsewhere.(11) The evaluation of the degree of difficulty uses two dimensions, and that of the nature of the effort combines three others. The thresholds used to define the classifications were fixed after various tests which make sure that the CDCO deals correctly with the most familiar and least ambiguous titles, reproducing in every point what could be called the unanimous judgement of writers on the subject. Moreover, applied to ambiguous and controversial titles, it allows a uniform and coherent decision to be reached. There is certainly nothing to guarantee that every one of these decisions, considered individually, will be the best possible one. On the other hand, we can be sure that the decisions as a whole are subject at the same rules, free of contradictions and unquestionably of a superior quality to the results of the often intuitive and improvised judgement of the researcher.

The real problem lies elsewhere: what are the dangers of applying descriptions and evaluations based on recent data to occupations of an earlier age?

4. THE PROBLEM OF LONGITUDINAL STUDIES

This difficulty is a major one. As a result of decades of far-reaching changes in the workplace, we must suspect that the content of many tasks has changed considerably, to the point of creating differences of degree and of nature which would be concealed by an anachronistic classification. In the first part of this paper, we examined the case of a multiplicity of titles referring to the same content. Here, we face the opposite problem: differing contents under the same title.

There are two ways of avoiding this problem. One consists of relying upon evaluations made at a time more or less contemporaneous with the period studied.(12) But no such evaluations exist for the Saguenay region. Moreover, the problem is further complicated because the period of our study runs from 1842 up until a few years ago. The second solution, an extremely rash one, consists of decreing a priori that no significant technological changes have taken place during the period studied.(13) Since we did not wish to take this way, two solutions remained: we could either construct a variable classification system geared to the different stages of economic evolution, or stick to the classification derived from the CDCO, and in that case, of course, anachronisms would have to be strictly controlled. We chose this last way, risky as it might seem at first glance. All our tests in this area seem to prove us right. They were intended to establish whether or not classifications made with the help of the CDCO on old occupations took account of the real content of these tasks, with respect to the evolution which has occurred since the middle of the nineteenth century. Two tests were carried out.(14) First, we compared the CDCO with well-known classifications, developed for old data and accepted as reliable: the purpose of this exercise was to measure the size and nature of the differences revealed.

Five classification grids were studied, involving british, american and canadian data.(15) From these works, a total sample of 657 different occupations were classified according to the terms of the CDCO. We then compared this classification with those of the authors, with particular attention to discordances introduced by the CDCO. These comparisons revealed that:

a) the next index of disagreement between the authors and the CDCO was about 1%. Furthermore, in these cases of disagreement, it was not evident that the authors' classifications were better than those of the CDCO.

b) the index of disagreement among the authors themselves was three times greater than in the case of the above (see Tables 3 and 4).

In a second test, with samples involving only Saguenay data, classifications produced a few years ago by SOREP researchers (cf. above), before the present method was adopted, were also compared to CDCO classifications. A third and final test focussed on the occupations of workers in a major industry, for the years 1926 - 1939. The Alcan company's aluminium manufacturing plant was chosen for this test. A survey-group of 140 occupations was classified by specialists from the company who were instructed to refer to the period indicated; that same group was then classified according to the terms of the CDCO. These results corroborated the previous results.(16)

The effect of anachronism which tests were intended to measure appears insignificant. Is it possible that, a few cases apart, we had somewhat exaggerated this difficulty? Actually, the crucial consideration is the margin of error introduced into the data-base as a whole, and it may happen that the classification procedures in current use by researchers are more to be feared. Our results suggest that, beyond its many imprecisions and confusions, occupational language posesses some sort of mechanism for adjustment and defence, whereby new contents (the results of the devision of tasks) or radical changes of content (e.g. those caused by rapid mechanization) tend to bring about the coining of new titles.

Whatever the truth of this tentative statement, the preceding data lead us to believe that a collection like the CDCO can be quite an adequate vehicle for the classification of data from an earlier time, and this methodological result is worthwhile, in itself.

TABLE 3

LIST OF OCCUPATIONS ON WHICH THE AUTHORS DISAGREE
(LEVEL OF DIFFICULTY)

N.B.: H = High Level L = Low level

Dashes indicate that the occupation was not present or not classifiable.

OCCUPATIONS	CLASSIFICATIONS				
	CDCO	S. Thernstrom	A. Armstrong	T. Hershberg	M. Katz
Accountant	H	L	H	L	–
Taxi Driver (or car)	L	L	H	L	–
Clerk	L	L	H	L	–
Cook	L	L	H	H	–
Cooper	L	L	H	H	H
Fireman	L	L	H	L	–
Barber	L	L	H	H	H
Bus Driver	L	L	H	–	–
Policeman	L	L	H	L	–
Sailor	L	L	H	L	–
Waiter	L	L	H	L	L
Guard (watchman)	L	L	H	L	–
Gardener	L	L	L	H	H
Pharmacist (apothecary)	H	H	–	L	H
Architect	H	H	H	L	–
Broker	H	H	H	L	–
Chemist	H	H	H	L	–
Editor	H	H	–	L	–
Florist	H	–	H	L	–
Inspector	H	–	H	L	–
Lumberer (woodcutter, lumberman, woodsman)	L	L	L	–	H
TOTAL: 21					

(SOREP) Saguenay Research Project

TABLE 4

LIST OF OCCUPATIONS ON WHICH THE AUTHORS DISAGREE
(NATURE OF THE EFFORT)

N.B.: M = Manual NM = Non-manual

Dashes indicate that the occupation was not present or not classifiable.

OCCUPATIONS	CLASSIFICATIONS				
	CDCO	S. Thernstrom	A. Armstrong	T. Hershberg	M. Katz
Baggageman (porter)	M	NM	M	M	–
Messenger	M	NM	M	–	–
Railroad conductor (and assimilated)	M	NM	–	M	–
Photographer	M	NM	–	M	–
Policeman (constable)	M	M	NM	M	NM
Farmer	M	–	M	NM	NM
Designer (draftman)	NM	NM	–	M	–
Civil Engineer*	NM	NM	M	M	–
TOTAL: 8					

*This is not to be confused with "engineer", meaning "millwright".

(SOREP) Saguenay Research Project

5. OCCUPATIONAL CATEGORIES AND SOCIAL CLASSES

The question of the use of the categories in the study of the classes is very complex. Accordingly, we offer not a new interpretation of this controversial problem, but simply a few remarks designed to situate our procedures and delineate their limits.

Initially, discussion of the relationships between categories and classes is doomed to confusion unless terms are exactly defined. The concept of class, as it is currently used, conceals radically different theoretical contents and analytical procedures. We have identified three of these. First, class is often confused with stratum of social rank and it is defined with reference to the elementary hierarchies created by the unequal distribution of attributes like income, prestige, education, etc. Here, we are dealing with the most descriptive and empirical use of the concept. In the second case, it can be defined with reference to the dynamic, to the actions of a social body. Then, the crucial matters or features are ideologies, class consciousness and class struggle, social movements insofar as they involve groups of actors in conflict, acting out the scenario of social change. Finally, class is also defined in terms of the social structure itself, whenever it relates to the fundamental structural divisions of society, or at least, to what these are perceived to be. With marxism, for example, these divisions concern the system of production, and the various positions recognized within the system mark out the places of the social classes. In these three senses, therefore, the classes correspond to either basic empirical entities, or actors in conflict, or fundamental structural positions. We need hardly to point out that, in reality, these three areas almost never overlap, giving us three very specific procedures among which to choose.

Let us return to the question of whether categories can reveal classes, or: whether the universe of the technical organization of tasks, on which the categories are based, is reducible to and continuous from that of the social division of labour, or concurrent with it and irreducible. In the first case, the categories can be used to identify the classes. For example, if the class is defined by wealth, it is possible to organize the categories into relevant entities. But if it is defined by its actions, the categories can become inoperative. For example, among the skilled workers, there are very militant, organized and active individuals along with other who are amorphous and, for all sorts of reasons, impossible to mobilize.(17) We can see that the categories cannot serve all the scientific endeavours aimed at depicting the "classes". If the first meaning of the concept is adopted, there is a high degree of compatibility between the two dimensions. With the other two meanings, this compatibility diminishes and the two dimensions sometimes become parallel and irreducible. Even there, however, the categories remain useful if not essential, as an aid to the identification of the social individuals who make up each of the classes (see above).(18) It should be clear that the researcher must make theoretical choices; that is, accept certain limitations. As the managers of a population register designed for many and various uses, we wished to avoid subjecting our grid to too specific or exclusive a theoretical orientation. And as historians, we had a marked preference for a very loosely-defined grid, as near as possible to empirical reality. In this way, we believe we have created an instrument capable of clarifying some important aspects of the phenomenon of the social classes. Finally, as we said, it is still always possible for the user to use a different method by returning to the raw data.

6. CLASSIFICATION BY SECTORS OF ACTIVITY

All of the above relates to the social component of the occupation. But occupation is also, as we know, an important economic indicator, in that it reveals the work structure or the distribution of labour force in the large sectors of the economy. This second kind of classification of occupational data is also based on a grid or on a form of nomenclature. Here, we used a procedure analogous to that described for the categories, that is: a critical appreciation of existing methodology and the construction of a new grid.

A. A critical review

We examined about 15 nomenclatures. Interestingly, they raised methodological problems similar to those encountered over the categories. The main problem, to which we will limit this discussion, lies in the confusion between nomenclature (that is to say, the classification by functions) and hierarchy.(19) Most grids divide the occupations and job-sectors into three parts, the primary, secondary and tiertiary, with the occasional addition of the quaternary. In reality, we are dealing with a special kind of hierarchy, which leads to ill-defined and very mixed divisions.

A new Zealand economist, A. G. B. Fisher, writing in 1935, seems to have been one of the first to define these areas of economic activity. This author already speaks of the wellknown trinity of primary, secondary and tertiary activities, which found very early use in the official statistics of New Zealand and Australia. Primary activities included agriculture, hunting, stock-breeding, forestry and mining. Secondary activities were part of industry, while the tertiary sector was sometimes defined by contrast to the above, and sometimes by reference to certain more sophisticated processing activities, like printing and publishing. This last ambiguity reveals the concern of Fisher and his successors. The noble sectors of the economy, in which the most interest is shown, are those susceptible to spectacular technological changes and which produce economic progress, such as industrial manufacture and the production of raw materials.

Obviously, this approach makes the third sector a sort of catch-all. This conception appears again, to a large extent, with Colin Clark (20), who only gives clear definition to secondary or manufacturing activities. For Clark, these activities consist of "the continuous large-scale conversion of raw materials into transportabel products" (p. 153). Curiously enough, this definition relegates to "the service industries" all activities connected with construction and transportation, as well as all of the "processes related to dressmaking, shoe-repair, etc." (p. 153). Clark quite explicitly reserves the secondary sector for industries with a high potential for mechanization and he makes it his indicator of technological development. The term "service industries" also includes all activities based on small-scale production of goods and all sorts of activities which require little equipment, with the exception of the railways, the merchant marine, the telephone, etc. (p. 207).

It is clear that, of the three "sectors" used, only the secondary sector is clearly defined, while the other two float within shifting limits. The sector of activities described by the heading "service" is particularaly heterogenous; there, administration and the arts are grouped with small manufacturing, the construction industry and transportation. This last sector is, in fact, deprived of all significance. Obviously, this instrument cannot be used to classify occupations according to sectors of

economic activity, which is not surprising, since that was not the author's purpose. Clark rather wished to construct an indicator which would allow him to compare, on an international level, the performance of certain leading industries with a high potential for mechanization and growth. To this end, he isolated them for observation and grouped them in a division which the author calls secondary.

Attempts made by A. Sauvy (1949, 1956), J. Fourastie (21) and several more recent authors have done little to oversome this problem of definitions, especially in the third sector. We therefore decided to ignore these major divisions, which are not really sectors, but groups of sectors based on one or more criteria, in the same way that we indicated that the classes can be groups of categories.

Also, it should be noted that our study revealed other, familiar problems, like the recourse to very irregular subsets, or to divisions which are not sectors but categories or classes (Journeymen, Proprietors, Clergymen, etc.). Considerable difficulties also arose from government censuses, whose divisions and definitions are modified so often.

B. A new nomenclature

The instrument we developed aims at maximum flexibility by avoiding all a priori grouping of sectors, by identifying the latter with the whole of the working activities rather than restricting to production as such, and by resorting to very loose sub-groups which can be assembled into a large number of groupings later.(22) The nomenclature consists of the following eighteen sectors:

1. Production raw materials and energy
2. Production of manufactured or semi-manufactored goods
3. Construction
4. Repairs, maintenance and similar activities
5. Transport, storage
6. Communications
7. Sales (wholesale)
8. Sales (Retail)
9. Sales (Unknown)
10. Finance
11. Special assistance to business and individuals
12. Civic protection, functioning of the State and connected services
13. Medical care, public health
14. Teaching, research
15. Religion
16. Social and community service
17. Recreation, leisure
18. Literary, artistic, cultural activities

Note that all this nomenclature and its divisions are based on the goal or the orientation of the working units (establishments, institutions, workshops, offices). Beyond that, the definition of the different sectors - each with numerous sub-divisions - seeks to reflect the real importance of the non-industrial activities, both social and cultural, by adressing them directly rather than by contrasting them to others as if they were merely peripheral.

6. CONCLUSION

The methods and instruments described above are now in use by SOREP, and they have so far performed satisfactorily. They were constructed in a spirit of compromise between two sometimes concurrent constraints: on the one hand, the need to develop a methodology capable of serving as wide a range of theoretical aprroaches as possible; on the other hand, the need to design processing methods which would allow for the fragile, imprecise but yet rich nature of empirical data. In the context and current phase of our work, this second concern has taken on considerable importance, which seems to have been justified; good theories and models cannot be built from bad data. The approach has also been to try to emphasize the need for a methodology at the same time explicit, transparent and concerned with coherence, which is an important condition of comparison and scientific dialogue.

More specifically, this research has singled out three points of method:

- Careful examination of occupational data (identification of synonyms, near-synonyms, etc.) is an essential preliminary to all analysis.

- Because of the confusion of titles and the often imprecise nature of the content of the jobs (23), it is wise to use the occupational data at the level of the categories and in the framework of career histories.

- The use of categories raises two problems: a) the criteria used in constructing them, and b) the evaluation procedure on which the classification of the occupations is based.

FOOTNOTES

* We wish to thank the Fonds FCAC (Québec) and the Université du Québec à Chicoutimi for their financial support to this research. We are also grateful to Christian Pouyez and Raymond Roy and all those who have contributed to it, directly or in other ways. In particular, Gordon Darroch, Paul-André Linteau and James Lehning were very helpful in criticizing an earlier draft of this paper, which has been presented at a Social Science History Association meeting in Washington (October 1983). The author, of course, remains fully responsible for weaknesses.

1 The skeptical reader may be tempted at this point to question the quality of the recording of the occupations in the certificates of the parish registers. However, sophisticated testing of our data (e.g. the simulation of "post-enumeration surveys") clears all doubts on this point. some of the tests are discussed in G. Bouchard, "L'analyse de la mobilité socio-professionnelle au Saguenay à l'aide de la reconstitution automatique des familles (An analysis of occupational mobility in the Saguenay region, using automatic family reconstruction) (upcoming publication).

2 All of the above confirms the results of a similar study undertaken as part of the Canadian social History Project (cf. Michael B. Katz, "Occupational classification in History", in Journal of Interdisciplinary History, Vol. 3, 1972, pp. 70 - 80). In this case, the analysis was performed by linking two name-lists, and consisted of comparing the declarations of occupations two by two. From this point of view, the files emerging from the family reconstruction allow us to go further, since they produce real itineraries of up to 25 entries, several of which are separated by very short intervals.

3 For an excellent insight into these problems, see Stuart Blumin, "The historical Study of Vertical Mobility", Historical Methods Newsletter, Vol. 1, No. 4, (Sept 1968), pp. 1 - 13; also, Clyde Griffin, "Occupatio-

nal Mobility in Nineteenth-Century America: Problems and Possibility", Journal of Social History, (Spring 1972), pp. 310 - 330.

4 We will restrict ourselves here to a very schematic description, omitting the illustrations and references appropriate to the modest explored. A more detailed discussion will be found in Working Paper No. 92 of the Saguenay Research Project. About forty grids were studied, from the productions of historians and sociologists whose work is generally known, if not celebrated.

5 We don't forget that the same appeal was made ten years ago by M. Katz, in the article quoted above (see footnote 2).

6 Note that the classification based on the work process also gives rise - as we shall see - to hierarchies. But these are not social as such.

7 Once again, we are forced to summarize and sacrifice necessary explanations, which will be found in Working Paper No. 92, already mentioned.

8 Taking the degree of difficulty of specialization as an example, we note that many grids use the categories: specialized, semi-specialized and non-specialized. But this refinement is often an illusion because it usually conceals very summary evaluation procedured, based on arbitrary judgements. All those who have experience of this work know what it implies.

9 The procedure is described in Working Paper No. 92 of the Saguenay Research Project.

10 For instance, category no. 20 on the grid contains all the occupations whose effort is predominantly manual and whose degree of difficulty is high, but for which the degree of independence is unknown. It should be noted that this indetermination is not final, since various other steps can be taken later to produce an exact evaluation (see, for example, Thomas Smith, "Reconstructing Occupational Structures: The Case of the Ambiguous Artisans", Historical Methods Newsletter, Vol. 8, No. 3 (June), pp. 134 - 146.

11 See Working Paper No. 71 of the Saguenay Research Project.

12 In the United States, researchers have fequently used a classification system developed by A. M. Edwards, "A Social Economic Grouping of the Gainful Workers of the United States", Journal of the American Statistical Association, vol. 27 (1933), pp. 377 - 387. In Great Britain, W. A. Armstrong used the Registrar-General's grid (W. A. Armstrong, "The Use of Information About Occupation", in E. A. Wrigley (ed.), Nineteenth-Century Society, Cambridge University Press, Cambridge, 1972, 448 pages; cf. pp. 215 - 223).

13 Somewhat surprisingly, this is done by Stephan Thernstrom, The Other Bostonians, Harvard University Press, 1973, 345 pages; (cf. pp. 292 - 293); who devotes scarcely a paragraph to this question in his study on the population of Boston.

14 It should be recalled that these tests related to only two of the criteria on our grid, the degree of difficulty and the nature of the effort.

15 They are: one by A. M. Edwards, as adapted by Stephan Thernstrom, The Other Bostonians, Harvard University Press, 1973, 345 pages; one by the Registrar-General (census service) of Great Britain, as presented by W. A. Armstron, "The use of Information About Occupation", in E. A. Wrigley (ed.), Nineteenth-Century Society, Cambridge University Press, Cambridge, 1972, 448 pages (cf. pp. 215 - 223); the Philadelphia Social History Project (Theodore Hershberg et Robert Dockhorn, "Occupational Classification", Historical Methods Newsletter, Vol. 9, Nos. 2 - 3 (mars - juin 1976), pp. 59 - 88); the Canadian Social History Project (Michael B. Katz, "Occupational Classification in History", in Journal of Interdisciplinary History, vol. 3, 1972, pp. 63 - 88); and finally, one constructed by a group of researchers attached to Quebec's Department of social Affairs.

16 A description of these tests and a detailed account of their results can be found in G. Bouchard, "L'utilisation des données socio-professionneles en histoire: le problème de la diachronie/The use of occupational data in history: the problem of longitudinal analysis", to appear in Histoire sociale/Social History (Ottawa), in summer of 1984.

17 This general problem is very well presented by Eric Olin Wright et David Hachen, "The American Class Structure", American sociaolgical Review, 1982, Vol. 47 (December 1982), pp. 709 - 726). See also Ronald Breiger, "The social class structure of occupational mobility". American Journal of Sociology, 1981, pp. 578 - 611.

18 For example: Michael B. Katz, Mark Stern, "Fertility, Class and Industrial Capitalism: Erie County, New-York, 1855 - 1915", American Quaterly, 33 (spring 1981), pp. 63 - 92; Eric Olin Wright, "Varieties of Marxist Conceptions of Class Structure", Politics and Society, 9, (1980), pp. 323 - 370; "Class and Occupation", Theory and Society, 9 (1980), pp. 177 - 214; J.-P. Courtheoux, "La structure en classes d'une population active. Hypothèses empiriques sur les catégories socio-professionnelles d'après les recensements francais de 1964 et 1962", Revue économique, vol. 16, No. 2, mars 1965, p. 246 - 275.

19 A detailed description of this critical review can be found in Working Paper No. 46 of the Saguenay Research Project.

20 Colin Clark, Les conditions du progrès économique, Paris, P. U. F., 1960.

21 Alfred Sauvy, "Progrès technique et répartition professionnelle de la population", Population, Vol. 4, No. 2 (janvier-mars 1949), pp. 57 - 76; Théorie générale de la population, Paris, 1956, tome I; Jean Fourastie, Le grand espoir du XXe siècle, Paris, Gallimard, 1963.

22 For a detailed description, see G. Bouchard, C. Pouyez, R. Roy, cf. "Le classement des professions par secteurs d'activités: apercu critique et présentation d'une nouvelle grille", Actualité économique (octobre-décembre 1979), pp. 585 - 605.

23 For example: S. B., living in a small village in the Saguenay region at the beginning of the 20th century, is a general merchant who runs an inn, farms a little and trades livestock. He acts as a hunting guide to Americans in the autumn and sometimes goes lumbering in the winter. Also, on two occasions, he states his occupation as bailiff.